

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.
OM protein - protein search, using sw model
Run on: November 21, 2005, 12:22:28 ; Search time 229 seconds
(without alignments)
742.499 Million cell updates/sec

Title: US-10-618-281-11
Perfect score: 1277
Sequence: 1 SLHLSRADWQYSORELDAV.....ELVAYQLERLQFIHELPS 241

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1266.5	99.2	242	Q9NPM1_HUMAN	Q9NPM1 homo sapien
2	1266.5	99.2	313	Q8VCV1_MOUSE	Q8VCV1 mus musculus
3	1266.5	99.2	313	Q6PCB6_HUMAN	Q6PCB6 homo sapien
4	1255.5	98.3	310	Q5ZJX1_CHICK	Q5ZJX1 gallus gall
5	1202.5	94.2	311	Q6DD70_XENLA	Q6DD70 xenopus lae
6	1199.5	93.9	310	Q6GL10_XENTR	Q6GL10 xenopus tro
7	1175.5	92.1	294	Q7ZV27_BRARE	Q7ZV27 brachydanio
8	1110.5	87.0	310	Q5X1J5_RAT	Q5X1J5 rattus norv
9	1110.5	87.0	310	Q9JW11_MOUSE	Q9JW11 mus musculus
10	1100.5	86.2	310	Q96G86_HUMAN	Q96G86 homo sapien
11	1095.5	85.8	310	Q6PJU2_HUMAN	Q6PJU2 homo sapien
12	1091.5	85.5	288	Q6DC55_XENLA	Q6DC55 xenopus lae
13	1091.5	85.4	317	Q4SD49_TETNG	Q4SD49 tetraodon n
14	1090.5	85.4	288	Q5ZJ01_CHICK	Q5ZJ01 gallus gall
15	1089.5	85.3	288	Q6DEY3_XENTR	Q6DEY3 xenopus tro
16	1087.5	85.2	288	Q5VST6_HUMAN	Q5VST6 homo sapien
17	1087.5	85.2	288	Q81Y03_HUMAN	Q81Y03 homo sapien
18	1086.5	85.1	288	Q7M759_MOUSE	Q7M759 mus musculus
19	1086.5	85.1	288	Q6AV17_RAT	Q6AV17 rattus norv
20	1086.5	85.0	293	Q5VST7_HUMAN	Q5VST7 homo sapien
21	1085.5	85.0	288	Q6DDV9_XENLA	Q6DDV9 xenopus lae
22	1083.5	84.8	288	Q86YB6_HUMAN	Q86YB6 homo sapien
23	1082.5	84.7	293	Q9Y377_HUMAN	Q9Y377 homo sapien
24	1066.5	83.5	310	Q5RGM9_HUMAN	Q5RGM9 homo sapien
25	1065.5	83.4	361	Q8WU99_HUMAN	Q8WU99 homo sapien
26	1062.5	83.2	305	Q5M904_XENTR	Q5M904 xenopus tro
27	1056.5	82.7	350	Q4RV21_TETNG	Q4RV21 tetraodon n
28	1052.5	80.8	288	Q7QDS9_ANOGA	Q7QDS9 anopheles g
29	1028.5	80.5	286	Q7PQX3_ANOGA	Q7PQX3 anopheles g
30	954	74.7	286	Q9VBX8_DROME	Q9VBX8 drosophila
31	837.5	65.6	183	Q58EV9_MOUSE	Q58EV9 mus musculus

32	825.5	64.6	332	2	Q21221_CABEL	Q21221 caenorhabdi
33	803.5	62.9	333	2	Q6IN34_CABBR	Q6IN34 caenorhabdi
34	769	60.2	336	2	Q9BWL0_HUMAN	Q9BWL0 homo sapien
35	742	58.1	223	2	Q4SCM5_TETNG	Q4SCM5 tetraodon n
36	741.5	58.1	274	2	Q5DU54_MOUSE	Q5DU54 mus musculus
37	677	53.0	244	2	Q9VBX9_DROME	Q9VBX9 drosophila
38	597	45.8	389	2	Q67VK9_ORYSA	Q67VK9 oryza sativ
39	595.5	46.6	389	2	Q6YU00_ORYSA	Q6YU00 oryza sativ
40	593.5	46.5	364	2	Q5N7J6_ORYSA	Q5N7J6 oryza sativ
41	586.5	45.9	365	2	Q9SB70_ARATH	Q9SB70 arabidopsis
42	584.5	45.8	336	2	Q9FF34_ARATH	Q9FF34 arabidopsis
43	573.5	44.9	399	2	Q9LI62_ARATH	Q9LI62 arabidopsis
44	572.5	44.8	361	2	Q9S7U7_ARATH	Q9S7U7 arabidopsis
45	571.5	44.8	369	2	Q9LY90_ARATH	Q9LY90 arabidopsis

ALIGNMENTS

RESULT 1
Q9NPM1_HUMAN
ID Q9NPM1_HUMAN PRELIMINARY; PRT; 242 AA.
AC Q9NPM1;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN NUCLEOTIDE SEQUENCE.
RP NUCLEOTIDE SEQUENCE.
RA Auffray C., Ansorge W., Ballabio A., Estivill X., Gibson K.,
RA Lehrach H., Poustka A., Lundeberg J.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RA Carim L., Estivill X., Escarceller M., Sumoy L.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
EMBL; AL390079; CAB98203.1; -, mRNA.
DR MEROPS; S09.053; -.
DR Ensembl; ENSG00000136379; Homo sapiens.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR InterPro; IPR000379; Ser_estrs.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 242 AA; 27190 MW; 3B9F1688F6D7A6A1 CRC64;

Query Match	99.2%	Score	1266.5	DB 2	Length	242			
Best Local Similarity	99.6%	Pred. No.	3.5e-107						
Matches	241	Conservative	0	Mismatches	0	Indels	1	Gaps	1
QY	1	SLHLSRADWQYSORELDAVEVFFSRTARDNRIGCMFVRCAPSSRYTLFLSHGNVDLQG	60						
Db	1	SLHLSRADWQYSORELDAVEVFFSRTARDNRIGCMFVRCAPSSRYTLFLSHGNVDLQG	60						
QY	61	MCSFYIGLSRINCINFSYDYSYGVSSGKPSKKNLYADIDAAQALRTYGVSPENIIL	120						
Db	61	MCSFYIGLSRINCINFSYDYSYGVSSGKPSKKNLYADIDAAQALRTYGVSPENIIL	120						
QY	121	YGQSIGTVPVDLASRYECAAATLHSPMSGRLVAFPDTRKTYCFDAFPIDIKISKVTSP	180						
Db	121	YGQSIGTVPVDLASRYECAAATLHSPMSGRLVAFPDTRKTYCFDAFPIDIKISKVTSP	180						
QY	181	VLVIHGTEDVIDFSLHGLAMYERCPRAVEPLWVEGAGHNDIELYAYQLERLQKQFI-HELP	239						
Db	181	VLVIHGTEDVIDFSLHGLAMYERCPRAVEPLWVEGAGHNDIELYAYQLERLQKQFISHHELP	240						
QY	240	NS 241							
Db	241	NS 242							

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2005, 12:34:03 ; Search time 164 Seconds
(without alignments)
614.005 Million cell updates/sec

Title: US-10-618-281-11

Perfect score: 1277

Sequence: 1 SLHLSEADWQYSORELDAV.....ELYAQYLERLQKFIHELPS 241

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgm2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgm2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 3: /cgm2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 4: /cgm2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 5: /cgm2_6/ptodata/1/pubpaa/US108_PUBCOMB.pep.*
- 6: /cgm2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1277	100.0	241	5	US-10-618-281-11
2	1155.5	90.5	308	4	US-10-312-354-4
3	1110.5	87.0	338	4	US-10-087-192-315
4	1100.5	86.2	310	3	US-09-934-392-2
5	1100.5	86.2	310	4	US-10-398-037-65
6	1095.5	85.8	310	4	US-10-115-479-6
7	1069.5	83.8	285	5	US-10-618-281-16
8	1066.5	83.5	290	4	US-10-115-479-4
9	1065	83.4	458	4	US-10-363-616-280
10	1061	83.1	305	5	US-10-618-281-21
11	1037.5	81.2	236	5	US-10-618-281-39
12	1026.5	80.4	403	5	US-10-450-763-58338
13	954	74.7	286	6	US-11-097-143-24690
14	774	60.6	236	4	US-10-087-192-318
15	665	52.1	244	6	US-11-097-143-32013
16	616.5	48.3	364	5	US-10-739-930-10670
17	607.5	47.6	367	4	US-10-425-115-345880
18	603.5	47.3	305	4	US-10-425-115-296034
19	597	46.8	389	4	US-10-437-963-111752
20	594.5	46.6	383	4	US-10-425-115-201722
21	594.5	46.6	452	4	US-10-425-114-60109
22	593.5	46.5	546	4	US-10-437-963-185558
23	591.5	46.3	344	5	US-10-450-763-48788
24	591.5	45.5	395	4	US-10-437-963-150508
25	578.5	45.3	415	4	US-10-425-114-67729
26	578.5	45.3	415	4	US-10-425-115-213872
27	572.5	44.8	367	4	US-10-424-599-197073

28	565.5	44.3	301	4	US-10-437-963-125073	Sequence 125073,
29	561	43.9	234	3	US-09-934-392-6	Sequence 6, Appli
30	560.5	43.9	350	4	US-10-425-114-70216	Sequence 70216, A
31	559.5	43.8	347	5	US-10-481-032A-624	Sequence 624, App
32	559.5	43.8	384	4	US-10-437-963-139334	Sequence 139334,
33	551.5	43.0	296	4	US-10-424-599-201993	Sequence 201993,
34	548.5	43.0	296	4	US-10-425-115-224377	Sequence 224377,
35	538	42.1	390	5	US-10-481-032A-376	Sequence 376, App
36	532	41.7	375	4	US-10-425-115-213029	Sequence 213029,
37	515.5	40.4	104	3	US-09-925-299-939	Sequence 939, App
38	515.5	40.4	104	3	US-09-925-299-939	Sequence 939, App
39	515.5	40.4	104	3	US-10-106-698-5792	Sequence 5792, Ap
40	486	38.1	96	5	US-10-450-763-41797	Sequence 41797, A
41	484	37.9	217	4	US-10-408-765A-1081	Sequence 1081, Ap
42	478	37.4	339	4	US-10-424-599-143301	Sequence 143301,
43	443.5	34.7	260	5	US-10-450-763-48787	Sequence 48787, A
44	409	32.0	227	5	US-10-739-930-7680	Sequence 7680, Ap
45	405	31.7	186	4	US-10-425-115-232924	Sequence 232924,

ALIGNMENTS

RESULT 1
US-10-618-281-11
; Sequence 11, Application US/10618281
; Publication No. US20040219609A1
; GENERAL INFORMATION:
; APPLICANT: Day, Anthony G.
; APPLICANT: Estell, David A.
; APPLICANT: Lyons, Eric H.
; APPLICANT: Yao, Jian
; TITLE OF INVENTION: Methods for Modulating Proteins Not
; TITLE OF INVENTION: Previously Known as Proteases
; FILE REFERENCE: GC773-2
; CURRENT APPLICATION NUMBER: US/10/618,281
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: US 60/395,325
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-618-281-11

Query Match	100.0%;	Score 1277;	DB 5;	Length 241;
Best Local Similarity	100.0%;	Pred. No. 1.2e-128;		
Matches 241;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	SLHLSEADWQYSORELDAVVEFFSRTARDNRNLGCMFVRCAPSSRYTLLFSHGNVDLQ	60	
DB	1	SLHLSEADWQYSORELDAVVEFFSRTARDNRNLGCMFVRCAPSSRYTLLFSHGNVDLQ	60	
QY	61	MCSFYIGLSRINCINIFSYDYGVSCKPSEKLYADIDAQALRTRYGVSPENIIL	120	
DB	61	MCSFYIGLSRINCINIFSYDYGVSCKPSEKLYADIDAQALRTRYGVSPENIIL	120	
QY	121	YGQSIGTPTVDLASRYECAAVILHSLPMSGLRVAFPPTRKTYCFDAPPSIDKISKVTS	180	
DB	121	YGQSIGTPTVDLASRYECAAVILHSLPMSGLRVAFPPTRKTYCFDAPPSIDKISKVTS	180	
QY	181	VLVIHGTDEVIDFSHGLMYERCPRAVEPLWVEGAGNDIELYAQYLERLQKFIHEL	240	
DB	181	VLVIHGTDEVIDFSHGLMYERCPRAVEPLWVEGAGNDIELYAQYLERLQKFIHEL	240	
QY	241	S 241		
DB	241	S 241		

RESULT 2

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2005, 12:07:52 ; Search time 188 Seconds
(without alignments)

563.246 Million cell updates/sec

Title: US-10-618-281-11

Perfect score: 1277

Sequence: 1 SLHLSRADWQVSQRELDV.....ELYAQYLERLKQFIHELPS 241

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_21.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1277	100.0	241	8 ADU24063	Adu24063 Human ser
2	1266.5	99.2	242	8 ADN05213	Adn05213 Antipsori
3	1266.5	99.2	329	8 ADK71002	Adk71002 Human est
4	1259.5	98.6	322	8 ADS10621	Ads10621 Human the
5	1259.5	98.6	329	8 ADS10620	Ads10620 Human the
6	1252.5	98.1	247	7 ADF60194	Adf60194 Human con
7	1252.5	98.1	247	8 ADS11892	Ads11892 Human the
8	1252.5	98.1	247	8 ADS11891	Ads11891 Human the
9	1155.5	90.5	308	5 AAU81978	Aau81978 Human sec
10	1110.5	87.0	338	7 ABM85300	Abm85300 Mouse pro
11	1100.5	86.2	310	4 AAM93226	Aam93226 Human pol
12	1100.5	86.2	310	5 AAU77137	Aau77137 Human alp
13	1100.5	86.2	310	5 ABP43541	Abp43541 Human sec
14	1100.5	86.2	310	8 ADL30608	Adl30608 Human pro
15	1095.5	85.8	310	7 ADE15976	Ade15976 G-coupled
16	1095.5	85.8	310	8 ADL33915	Adl33915 Human G-c
17	1086	85.0	293	8 ADL27199	Adl27199 Amino aci
18	1069.5	83.8	285	8 ADU24068	Adu24068 Human ser
19	1066.5	83.5	290	7 ADE15974	Ade15974 G-coupled
20	1066.5	83.5	290	8 ADL33913	Adl33913 Human G-c
21	1066.5	83.5	310	8 ABM80127	Abm80127 Tumour-as
22	1065	83.4	361	8 ADO20182	Ado20182 Human PRO
23	1065	83.4	361	8 ABM81992	Abm81992 Tumour-as
24	1065	83.4	458	5 ABP62843	Abp62843 Human pol

25	1061	83.1	305	8 ADU24073	Adu24073 Human ser
26	1037.5	81.2	236	8 ADU24091	Adu24091 Human ser
27	1026.5	80.4	403	4 ABG27979	Abg27979 Novel hum
28	954	74.7	286	4 ABM65966	Abm65966 Drosophil
29	774	60.6	236	7 ABM85301	Abm85301 Human pro
30	769	60.2	236	9 ADY17790	Ady17790 PRO polyp
31	665	52.1	244	4 ABM8407	Abm8407 Drosophil
32	616.5	48.3	364	8 ADT60593	Adt60593 Plant pol
33	594.5	46.6	452	8 ADY04294	Ady04294 Plant ful
34	591.5	46.3	344	4 ABG18429	Abg18429 Novel hum
35	586.5	45.9	331	3 AAG20746	Aag20746 Arabidops
36	586.5	45.9	358	3 AAG20745	Aag20745 Arabidops
37	586.5	45.9	365	3 AAG20744	Aag20744 Arabidops
38	586.5	45.9	652	3 AAG51304	Aag51304 Arabidops
39	586.5	45.9	679	3 AAG51303	Aag51303 Arabidops
40	586.5	45.9	686	3 AAG51302	Aag51302 Arabidops
41	586	45.9	190	7 ADF58699	Adf58699 Human pol
42	578.5	45.3	415	8 ADY11914	Ady11914 Plant ful
43	572.5	44.8	361	8 ADN73587	Adn73587 Thale cre
44	572.5	44.8	361	8 ADN73587	Adn73587 Thale cre
45	571.5	44.8	335	3 AAG14422	Aag14422 Arabidops

ALIGNMENTS

RESULT 1

ADU24063

ID ADU24063 standard; protein; 241 AA.

XX ADU24063;

DT 27-JAN-2005 (first entry)

DE Human serine peptidase enzyme #3.

XX Protease; cancer; immune-related disorder; cardiovascular disease;
KW brain-associated disease; neuronal-associated disease;
KW metabolic disorder; haematopoietic; cytostatic; immunosuppressive;
KW cardiovascular; cerebroprotective; neuroprotective; human;
KW serine peptidase; enzyme; peripheral nervous system; Alzheimer's disease;
KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
KW infection; ocular disease; migraine; pain; sexual dysfunction;
KW mood disorder; attention disorder; cognition disorder; hypotension;
KW hypertension; psychotic disorder; neurological disorder; dyskinesia;
KW organ transplant rejection.

XX Homo sapiens.

XX Key Location/Qualifiers
FH Domain 36..240
FT /note = Protease active domain
XX

US2004219609-A1.

04-NOV-2004.

11-JUL-2003; 2003US-00618281.

12-JUL-2002; 2002US-0395325P.

(DAYA/) DAY A G.

(ESTE/) ESTELL D A.

(LYON/) LYONS E H.

(YAO/) YAO J.

Day AG, Estell DA, Lyons EH, Yao J;

WPI; 2004-794441/78.

Identifying compound that modulates activity of protease, by contacting
protease with test compound, measuring activity of protease before and
after contacting step, and determining whether test compound modulates

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2005, 12:20:17 ; Search time 16 Seconds
(without alignments)
1449,265 Million cell updates/sec

Title: US-10-618-281-11
Perfect score: 1277
Sequence: 1 SLHLERADWQYSQRELDV.....ELYAQYLERLKQFIHELPS 241

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 segs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	825.5	64.6	405	2 T23321	hypothetical prote
2	586.5	45.9	365	2 T05558	hypothetical prote
3	571.5	44.8	369	2 T48612	hypothetical prote
4	545.5	42.7	341	2 B8269	F21F23.4 protei
5	534.5	41.9	422	2 D86446	hypothetical prote
6	525.5	41.2	316	2 C84635	hypothetical prote
7	481.5	37.7	256	2 G96692	hypothetical prote
8	462	36.2	307	2 C83363	hypothetical prote
9	399	31.2	245	2 AD2564	hypothetical prote
10	363.5	28.5	305	2 T20470	hypothetical prote
11	323.5	25.3	335	2 T20465	hypothetical prote
12	239.5	18.8	281	2 G70734	hypothetical prote
13	213.5	16.7	293	2 E65030	hypothetical prote
14	207.5	16.2	293	2 C85898	probable proteinase
15	207.5	16.2	293	2 H91053	probable enzyme li
16	202	15.8	284	2 S51294	probable membrane
17	189.5	14.8	292	2 AB0825	probable membrane
18	159.5	12.5	247	2 T17237	hypothetical prote
19	150	11.7	301	2 G83182	hypothetical prote
20	138	10.8	254	2 C95273	hypothetical prote
21	134.5	10.5	482	2 A44638	hypothetical prote
22	131.5	10.3	251	2 AF2674	hypothetical prote
23	131.5	10.3	4558	2 C82199	RTX toxin RtxA Vc1
24	122.5	9.6	261	2 T35708	hydrolase - Strept
25	122	9.6	238	2 C69334	2-hydroxy-6-oxonep
26	120.5	9.4	302	2 AB3196	arylester hydrolas
27	118.5	9.3	251	2 AD3177	dipeptidyl peptida
28	116	9.1	282	2 G75384	probable lipase -
29	116	9.1	318	2 T00552	lysophospholipase

probable hydrolase
lipase-like protei
acetoin dehydrogen
probable phosphol
hypothetical prote
hypothetical prote
triacylglycerol li
hydrolase, alpha/b
lipase-like protei
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
streptothricin-ace
hypothetical prote
lysophospholipase

ALIGNMENTS

RESULT 1

T23321
hypothetical protein K04G2.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change -09-Jul-2004
C:Accession: T23321
R:Gardner, A.
submitted to the EMBL Data Library, July 1996
A:Reference number: Z19727
A:Accession: T23321
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-405 <WIL>
A:Cross-references: UNIPROT:Q21221; UNIPARC:UPI000007914; EMBL:Z75712; PTDN:CAB00039.1;
A:Experimental source: clone K04G2
C:Genetics:
A:Gene: CESP.K04G2.2
A:Map position: 1
A:introns: 2/2; 31/1; 104/2; 355/1

Query Match 64.6%; Score 825.5; DB 2; Length 405;
Best Local Similarity 66.1%; Pred. No. 5.4e-66;
Matches 152; Conservative 36; Mismatches 41; Indels 1; Gaps 1;
QY 7 RADWQYSQREL-DAVEVFPSRTARDNRLGCMFVRCAPSSRYTLLFSGHNAVDLGQMSFY 65
DB 146 RAANPHQEVDMANCVEMRITRRNRNVACTMIRPLPNSHFLLFSGHNAVDLGQMTSFL 205
QY 66 IGLGRINCNIFSYDYGVSSEKPKSEKNLYADIDAAQALTRYGVSPENILYQSI 125
DB 206 YLGLPHLNCNVSFYDYGVSSEKPKSEKNLYADITAAPELKKSEFVGPKEKILYQSI 265
QY 126 GTVPVTLASRYECAAATLHSLPLMSGLRVAFDPTRKTYCFDAFPISIDKISKVTSPLVITH 185
DB 266 GTVPSVDLASHREDLAAVLHSLPLMSGLRVAFPGITTTTWCDDAFPSIEKVPKVCPTLVIH 325
QY 186 GTEDEVIDFSGHGLMYRCRAVEPLWVEGAGHNDIELYAQYLERLKQFI 235
DB 326 GTDDEVIDFSGHGVSIYRCPTSEVPLWVPGAGHNDVHAAVLERLRSFI 375

RESULT 2

T05558
hypothetical protein F22K18.40 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T05558
R:Bevan, M.; Wedler, H.; Wedler, E.; Wambutt, R.; Hoheisel, J.; Mayer, K.F.
submitted to the Protein Sequence Database, February 1999
A:Reference number: Z15419
A:Accession: T05558
A:Molecule type: DNA

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2005, 12:33:53 ; Search time 46 Seconds
(without alignments)
433.149 Million cell updates/sec

Title: US-10-618-281-11
Perfect score: 1277
Sequence: 1 SHLSERADWQYSORELDAV.....ELVAYLERLKQFIHELPS 241

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/prodata/1/iaa/5_COMB.pep.*
2: /cgn2_6/prodata/1/iaa/6_COMB.pep.*
3: /cgn2_6/prodata/1/iaa/H_COMB.pep.*
4: /cgn2_6/prodata/1/iaa/PCUS_COMB.pep.*
5: /cgn2_6/prodata/1/iaa/RG_COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	943.5	73.9	374	2	US-09-270-767-45545
2	286	22.4	152	2	US-09-270-767-61054
3	227	17.8	263	2	US-09-902-540-12703
4	184.5	14.4	275	2	US-09-248-796A-15244
5	150	11.7	661	2	US-09-252-991A-18225
6	149.5	11.7	312	2	US-09-107-532A-5684
7	126.5	9.9	2807	2	US-09-543-681A-4980
8	118.5	9.3	339	2	US-09-489-847-350
9	118.5	9.3	554	2	US-09-489-847-352
10	117.5	9.2	398	2	US-09-107-532A-6418
11	110	8.6	305	2	US-09-462-845-9
12	110	8.6	305	2	US-10-402-312-9
13	110	8.6	305	2	US-10-401-437-9
14	110	8.6	305	2	US-10-402-067-9
15	110	8.6	305	2	US-10-401-436-9
16	109.5	8.6	271	2	US-09-710-279-2822
17	109.5	8.6	271	2	US-09-710-279-3308
18	109.5	8.6	284	2	US-09-134-001C-2927
19	107	8.4	300	2	US-09-355-166-4
20	106.5	8.3	374	2	US-09-252-991A-27773
21	106	8.3	296	2	US-09-902-540-10861
22	100	7.8	346	1	US-08-602-359A-34
23	99.5	7.8	255	2	US-09-489-039A-9075
24	98.5	7.7	297	1	US-08-602-359A-37
25	95	7.4	298	2	US-09-902-540-10513
26	93.5	7.3	310	2	US-09-328-352-7313
27	93.5	7.3	589	2	US-09-252-991A-31105

28	7.3	210	2	US-09-902-540-11094	Sequence 11094, A
29	7.1	278	2	US-09-522-401-2	Sequence 2, Appli
30	7.0	342	2	US-09-489-039A-8276	Sequence 8276, Ap
31	7.0	285	2	US-09-328-352-7483	Sequence 7483, Ap
32	7.0	293	2	US-09-328-352-6002	Sequence 6002, A
33	88.5	651	2	US-09-902-540-11752	Sequence 11752, A
34	88	841	2	US-09-518-550-44	Sequence 44, Appl
35	87.5	369	2	US-09-198-452A-1001	Sequence 1001, Ap
36	87.5	372	2	US-09-438-185A-931	Sequence 931, App
37	87	261	2	US-09-902-540-13754	Sequence 13754, A
38	86	248	2	US-08-935-263-14	Sequence 14, Appl
39	86	248	2	US-09-594-185-14	Sequence 14, Appl
40	86	248	2	US-10-033-078-14	Sequence 14, Appl
41	86	248	2	US-10-763-933-14	Sequence 14, Appl
42	86	258	2	US-09-252-991A-20260	Sequence 20260, A
43	86	277	2	US-09-424-349A-6	Sequence 6, Appli
44	86	334	2	US-09-902-540-11426	Sequence 11426, A
45	85.5	319	2	US-09-328-352-5110	Sequence 5110, Ap

ALIGNMENTS

RESULT 1
US-09-270-767-45545
; Sequence 45545, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 45545
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-45545

Query Match		73.9%	Score 943.5;	DB 2;	Length 374;
Best Local Similarity		72.0%	Pred. No. 2.4e-97;		
Matches 170;		Conservative 37;	Mismatches 28;	Indels 1;	Gaps 1;
QY	1	SHLSERADWQYSORELDAVEVFFSRTARDNRLGCMFVRCAPSSRYTLFLFSHGNAVDLGG	60		
Db	54	NLQLFDRAEWQYSEREKSKVEAFFRTSRGNLITCIYVRCNKAKYTLFLFSHGNAVDLGG	113		
QY	61	MCSFYIGLSRINCNIIFYSYSGYVSSGKPKSEKNLYADIDAQAL-RTRYGVSPENII	119		
Db	114	MSSFYTLGSQLNCINIFGYDYSYGVMGSKPKSEKNLYADIEAAQAMXRTFNISPETII	173		
QY	120	LYGQSIGVPTVDLASRYECARVILHSLPLMSGLRVAFDPDTRKTYCFDAPSDKISKYMS	179		
Db	174	LYGQSIGVPTVDLASRHEVGAVILHSLPLMSGLRVVFNRTKRTWTFDAPSDKIAKVA	233		
QY	180	PVLVIHGTEDEVIDFSHGLAMVRCPRAVEPLWVEGAGHNDIELVAYLERLKQFI	235		
Db	234	PVLVIHGTDDEVIDFSHGIYERCPKTVPEPPWVEGAGHNDVELHPHYERLKF	289		

RESULT 2
US-09-270-767-61054
; Sequence 61054, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2005, 12:37:13 ; Search time 6 seconds
(without alignments)
45.380 Million cell updates/sec

Title: US-10-618-281-11

Perfect score: 1277

Sequence: 1 SLHBRADWQSQRELDV.....ELYAQYLERLKQFIHELPSN 241

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 8323 seqs, 1129788 residues

Total number of hits satisfying chosen parameters: 8323

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA New:

- 1: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	107	8.4	300	7	US-11-179-977-4
2	78	6.1	256	7	US-11-179-977-2
3	74	5.8	872	7	US-11-077-550-145
4	69.5	5.4	319	7	US-11-109-156-38
5	69.5	5.4	323	7	US-11-109-156-37
6	68.5	5.4	242	7	US-11-179-977-14
7	65.5	5.1	330	1	US-10-689-742-82
8	65.5	5.1	657	7	US-11-179-977-1
9	64	5.0	289	7	US-11-179-977-11
10	63.5	5.0	325	7	US-11-074-176-370
11	63.5	5.0	429	7	US-11-074-176-248
12	63	4.9	521	7	US-11-109-156-34
13	60	4.7	298	7	US-11-179-977-6
14	59.5	4.7	520	7	US-11-098-662-12
15	59.5	4.7	799	7	US-11-074-176-348
16	59.5	4.7	805	7	US-11-074-176-172
17	59	4.6	119	1	US-10-502-145-25
18	59	4.6	256	7	US-11-179-977-15
19	59	4.6	318	7	US-11-179-977-16
20	58.5	4.6	211	7	US-11-098-662-16
21	58.5	4.6	2376	7	US-11-098-662-14
22	58.5	4.6	2715	7	US-11-096-051-2
23	58	4.5	2280	7	US-11-022-562-211
24	57.5	4.5	876	7	US-11-077-550-82
25	57.5	4.5	1332	7	US-11-091-643-18

26	57.5	4.5	2721	7	US-11-096-051-10	Sequence 10, Appl
27	57.5	4.5	2725	7	US-11-096-051-8	Sequence 8, Appl
28	57	4.5	390	7	US-11-082-389-42	Sequence 42, Appl
29	57	4.5	390	7	US-11-082-389-44	Sequence 44, Appl
30	57	4.5	569	1	US-10-632-150-2	Sequence 2, Appl
31	57	4.5	902	7	US-11-057-058-64	Sequence 64, Appl
32	56.5	4.4	863	7	US-11-077-550-36	Sequence 36, Appl
33	56.5	4.4	876	7	US-11-077-550-106	Sequence 106, App
34	56.5	4.4	876	7	US-11-077-550-108	Sequence 108, App
35	56.5	4.4	910	1	US-10-131-826A-112	Sequence 112, App
36	56.5	4.4	932	7	US-11-017-550-65	Sequence 65, Appl
37	56.5	4.4	1127	7	US-11-077-550-54	Sequence 54, Appl
38	56.5	4.4	1129	7	US-11-077-550-52	Sequence 52, Appl
39	56	4.4	174	1	US-10-984-376-7	Sequence 7, Appl
40	56	4.4	467	7	US-11-082-389-334	Sequence 334, App
41	56	4.4	475	1	US-10-510-386-98	Sequence 98, Appl
42	56	4.4	552	1	US-10-131-826A-196	Sequence 196, App
43	55.5	4.3	406	1	US-10-131-826A-66	Sequence 66, Appl
44	55.5	4.3	406	1	US-10-131-826A-258	Sequence 258, App
45	55.5	4.3	487	1	US-10-131-826A-528	Sequence 528, App

ALIGNMENTS

RESULT 1
US-11-179-977-4
; Sequence 4, Application US/11179977
; Publication No. US20050249789A1
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Alpha/Beta Hydrolase-Fold Enzymes
; FILE REFERENCE: GC511-PCT
; CURRENT APPLICATION NUMBER: US/11/179,977
; CURRENT FILING DATE: 2005-07-12
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Bacillus
US-11-179-977-4

Query Match 8.4%; Score 107; DB 7; Length 300;
Best Local Similarity 21.7%; Pred. No. 2.9e-05;
Matches 49; Conservative 41; Mismatches 94; Indels 42; Gaps 11;

QY	47	TLFSGH---	NAVDLGOMCSFYIGLSR	INCNI	IFSYD	YSGYSG	KGPKSE	KNLYA	--DID	101																																	
Db	83	TLIICHGV	TMTNVLSLKY	NHFLDLG	---	WNVLIYD	HRRHG	SGGK	TKTSYGFYKDDLN	138																																	
QY	102	AAWQAL	TRYGVSPEN	II	LYG	SGIGT	VP	TV	-----	DLASRY--ECAA	VILH	SPLMS	150																														
Db	139	KVVSLL	KNK--	TNHR	GLIG	HGSG	MAV	TALLY	VAGAH	CSGD	ADFY	ADCP	FACF	DEQLAY	196																												
QY	151	GLRVA	-----	FP	TRK	TYC	FD	AP	SP	ID	KISK	VTSP	VLV	HG	TE	DE	VD	F	SH	196																							
Db	197	RLRAE	YRLP	SW	PLP	LP	AD	FL	KLR	GGY	RA	RE	VS	PL	AV	ID	KIE	KP	VL	FI	HS	KDD	DD	Y	P	V	SS	256															
QY	197	GLAM	YRC	PR	AV	E	PL	VE	GA	GH	ND	I	E	L	Y	A	Q	---	Y	L	E	R	L	K	Q	F	I	H	E	L	238												
Db	257	TEEL	Y	E	K	-	K	G	P	K	A	L	I	A	E	N	G	E	H	A	M	S	-	Y	T	K	N	R	H	T	Y	R	K	T	V	Q	E	F	L	D	N	M	300

RESULT 2
US-11-179-977-2
; Sequence 2, Application US/11179977
; Publication No. US20050249789A1
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Alpha/Beta Hydrolase-Fold Enzymes
; FILE REFERENCE: GC511-PCT
; CURRENT APPLICATION NUMBER: US/11/179,977